

1
SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: DOMPE' S.p.A.
(B) STREET: Via Campo di Pile
(C) CITY: L'AQUILA
(E) COUNTRY: ITALY
(F) POSTAL CODE (ZIP): 67100

(ii) TITLE OF INVENTION: RECOMBINANT PROTEINS DERIVED FROM HGF AND MSP

(iii) NUMBER OF SEQUENCES: 4

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1725 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATGTGGGTGA CCAAATCCT GCCAGCCCTG CTGCTGCAGC ATGTCCTCCT GCATCTCCTC 60
CTGCTCCCCA TCGCCATCCC CTATGCAGAG GGACAAAGGA AAAGAAGAAA TACAATTCAT 120
GAATTCAAAA AATCAGCAAA GACTACCCTA ATCAAAATAG ATCCAGCACT GAAGATAAAA 180
ACCAAAAAAG TGAATACTGC AGACCAATGT GCTAATAGAT GTACTAGGAA TAAAGGACTT 240
CCATTCACCTT GCAAGGCTTT TGTTTTTGAT AAAGCAAGAA AACAATGCCT CTGGTTCCCC 300
TTCAATAGCA TGTCAAGTGG AGTGAAAAAA GAATTTGGCC ATGAATTTGA CCTCTATGAA 360
AACAAAGACT ACATTAGAAA CTGCATCATT GGTAAGGAC GCAGCTACAA GGAACAGTA 420
TCTATCACTA AGAGTGGCAT CAAATGTCAG CCCTGGAGTT CCATGATACC ACACGAACAC 480
AGCTATCGGG GTAAAGACCT ACAGGAAAAC TACTGTCGAA ATCCTCGAGG GGAAGAAGGG 540
GGACCCTGCT GTTTCACAAG CAATCCAGAG GTACGCTACG AAGTCTGTGA CATTCTCAG 600

sub
B11
Cont. 1

TTGTCAGAAG TTGAATGCAT GACCTGCAAT GGGGAGAGTT ATCGAGGTCT CATGGATCAT 660
ACAGAATCAG GCAAGATTG TCAGCGCTGG GATCATCAGA CACCACACCG GCACAAATTC 720
TTGCCTGAAA GATATCCCGA CAAGGGCTTT GATGATAATT ATTGCCGCAA TCCCGATGGC 780
CAGCCGAGGC CATGGTGCTA TACTCTTGAC CCTCACACCC GCTGGGAGTA CTGTGCAATT 840
AAAACATGCG CTGACAAAGC TTCGGGCGGT GCGGTTCTG GTGGCGGTGG CTCCGGCGGT 900
GGCGGTTCTC TAGAGGGACA AAGGAAAAGA AGAAATACAA TTCATGAATT CAAAAAATCA 960
GCAAAGACTA CCCTAATCAA AATAGATCCA GCACTGAAGA TAAAAACCAA AAAAGTGAAT 1020
ACTGCAGACC AATGTGCTAA TAGATGTACT AGGAATAAAG GACTTCCATT CACTTGCAAG 1080
GCTTTTGTTT TTGATAAAGC AAGAAAACAA TGCCTCTGGT TCCCCTTCAA TAGCATGTCA 1140
AGTGGAGTGA AAAAAGAATT TGGCCATGAA TTTGACCTCT ATGAAAACAA AGACTACATT 1200
AGAAACTGCA TCATTGGTAA AGGACGCAGC TACAAGGGAA CAGTATCTAT CACTAAGAGT 1260
GGCATCAAAT GTCAGCCCTG GAGTTCCATG ATACCACACG AACACAGCTA TCGGGGTAAA 1320
GACCTACAGG AAAACTACTG TCGAAATCCT CGAGGGGAAG AAGGGGGACC CTGGTGTTTC 1380
ACAAGCAATC CAGAGGTACG CTACGAAGTC TGTGACATTC CTCAGTGTTT AGAAGTTGAA 1440
TGCATGACCT GCAATGGGGA GAGTTATCGA GGTCTCATGG ATCATAACAG ATCAGGCAAG 1500
ATTTGTCAGC GCTGGGATCA TCAGACACCA CACCGGCACA AATTCTTGCC TGAAAGATAT 1560
CCCGACAAGG GCTTTGATGA TAATTATTGC CGCAATCCCG ATGGCCAGCC GAGGCCATGG 1620
TGCTATACTC TTGACCCTCA CACCCGCTGG GACTACTGTG CAATTAAAAC ATGCGCTGAC 1680
AAAGCTGACG ACGACGACAA ACACCACCAC CACCACCACC ACTAG 1725

3

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 574 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Trp Val Thr Lys Leu Leu Pro Ala Leu Leu Leu Gln His Val Leu
 1 5 10 15

Leu His Leu Leu Leu Leu Pro Ile Ala Ile Pro Tyr Ala Glu Gly Gln
 20 25 30

Arg Lys Arg Arg Asn Thr Ile His Glu Phe Lys Lys Ser Ala Lys Thr
 35 40 45

Thr Leu Ile Lys Ile Asp Pro Ala Leu Lys Ile Lys Thr Lys Lys Val
 50 55 60

Asn Thr Ala Asp Gln Cys Ala Asn Arg Cys Thr Arg Asn Lys Gly Leu
 65 70 75 80

Pro Phe Thr Cys Lys Ala Phe Val Phe Asp Lys Ala Arg Lys Gln Cys
 85 90 95

Leu Trp Phe Pro Phe Asn Ser Met Ser Ser Gly Val Lys Lys Glu Phe
 100 105 110

Gly His Glu Phe Asp Leu Tyr Glu Asn Lys Asp Tyr Ile Arg Asn Cys
 115 120 125

Ile Ile Gly Lys Gly Arg Ser Tyr Lys Gly Thr Val Ser Ile Thr Lys
 130 135 140

Ser Gly Ile Lys Cys Gln Pro Trp Ser Ser Met Ile Pro His Glu His
 145 150 155 160

Ser Tyr Arg Gly Lys Asp Leu Gln Glu Asn Tyr Cys Arg Asn Pro Arg
 165 170 175

Gly Glu Glu Gly Gly Pro Trp Cys Phe Thr Ser Asn Pro Glu Val Arg
 180 185 190

Tyr Glu Val Cys Asp Ile Pro Gln Cys Ser Glu Val Glu Cys Met Thr
 195 200 205

Cys Asn Gly Glu Ser Tyr Arg Gly Leu Met Asp His Thr Glu Ser Gly
 210 215 220

Sub
 B11
 Cont'd

Sequence

Lys 225	Ile	Cys	Gln	Arg	Trp 230	Asp	His	Gln	Thr	Pro 235	His	Arg	His	Lys	Phe 240
Leu	Pro	Glu	Arg	Tyr 245	Pro	Asp	Lys	Gly	Phe 250	Asp	Asp	Asn	Tyr	Cys 255	Arg
Asn	Pro	Asp	Gly 260	Gln	Pro	Arg	Pro	Trp 265	Cys	Tyr	Thr	Leu	Asp 270	Pro	His
Thr	Arg	Trp 275	Glu	Tyr	Cys	Ala	Ile 280	Lys	Thr	Cys	Ala	Asp 285	Lys	Ala	Ser
Gly	Gly 290	Gly	Gly	Ser	Gly	Gly 295	Gly	Gly	Ser	Gly	Gly 300	Gly	Gly	Ser	Leu
Glu 305	Gly	Gln	Arg	Lys	Arg 310	Arg	Asn	Thr	Ile	His 315	Glu	Phe	Lys	Lys	Ser 320
Ala	Lys	Thr	Thr	Leu 325	Ile	Lys	Ile	Asp	Pro 330	Ala	Leu	Lys	Ile	Lys	Thr 335
Lys	Lys	Val	Asn 340	Thr	Ala	Asp	Gln	Cys 345	Ala	Asn	Arg	Cys	Thr 350	Arg	Asn
Lys	Gly	Leu 355	Pro	Phe	Thr	Cys	Lys	Ala	Phe	Val	Phe	Asp 365	Lys	Ala	Arg
Lys 370	Gln	Cys	Leu	Trp	Phe	Pro	Phe	Asn	Ser	Met	Ser 380	Ser	Gly	Val	Lys
Lys 385	Glu	Phe	Gly	His	Glu 390	Phe	Asp	Leu	Tyr	Glu 395	Asn	Lys	Asp	Tyr	Ile 400
Arg	Asn	Cys	Ile	Ile 405	Gly	Lys	Gly	Arg	Ser	Tyr	Lys	Gly	Thr	Val	Ser 415
Ile	Thr	Lys	Ser 420	Gly	Ile	Lys	Cys	Gln 425	Pro	Trp	Ser	Ser	Met	Ile	Pro 430
His	Glu	His 435	Ser	Tyr	Arg	Gly	Lys 440	Asp	Leu	Gln	Glu	Asn 445	Tyr	Cys	Arg
Asn 450	Pro	Arg	Gly	Glu	Glu	Gly 455	Gly	Pro	Trp	Cys	Phe 460	Thr	Ser	Asn	Pro
Glu 465	Val	Arg	Tyr	Glu	Val 470	Cys	Asp	Ile	Pro	Gln 475	Cys	Ser	Glu	Val	Glu 480
Cys	Met	Thr	Cys	Asn 485	Gly	Glu	Ser	Tyr	Arg	Gly	Leu	Met	Asp	His	Thr 495
Glu	Ser	Gly	Lys 500	Ile	Cys	Gln	Arg	Trp 505	Asp	His	Gln	Thr	Pro	His	Arg
His	Lys	Phe 515	Leu	Pro	Glu	Arg	Tyr 520	Pro	Asp	Lys	Gly	Phe 525	Asp	Asp	Asn

Lys Ala Asp Asp Asp Asp Lys His His His His His His His
565 570

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1692 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATGGGGTGGC	TCCCACTCCT	GCTGCTTCTG	ACTCAATGCT	TAGGGGTCCC	TGGGCAGCGC	60
TCGCCATTGA	ATGACTTCCA	AGTGCTCCGG	GGCACAGAGC	TACAGCACCT	GCTACATGCG	120
GTGGTGCCCG	GGCCTTGGCA	GGAGGATGTG	GCAGATGCTG	AAGAGTGTGC	TGGTCGCTGT	180
GGGCCCTTAA	TGGACTGCCG	GGCCTTCCAC	TACAACGTGA	GCAGCCATGG	TTGCCAACTG	240
CTGCCATGGA	CTCAACACTC	GCCCCACACG	AGGCTGCGGC	GTTCTGGGCG	CTGTGACCTC	300
TTCCAGAAGA	AAGACTACGT	ACGGACCTGC	ATCATGAACA	ATGGGGTTGG	GTACCGGGGC	360
ACCATGGCCA	CGACCGTGGG	TGGCCTGCCC	TGCCAGGCTT	GGAGCCACAA	GTTCCCGAAT	420
GATCACAAGT	ACACGCCCAC	TCTCCGGAAT	GGCCTGGAAG	AGAACTTCTG	CCGTAACCCT	480
GATGGCGACC	CCGGAGGTCC	TTGGTGCTAC	ACAACAGACC	CTGCTGTGCG	CTTCCAGAGC	540
TGCGGCATCA	AATCCTGCCG	GGAGGCCGCG	TGTGTCTGGT	GCAATGGCGA	GGAATACCGC	600
GGCGCGGTAG	ACCGCACGGA	GTCAGGGCGC	GAGTGCCAGC	GCTGGGATCT	TCAGCACCCG	660
CACCAGCACC	CCTTCGAGCC	GGGCAAGTTC	CTCGACCAAG	GTCTGGACGA	CAACTATTGC	720
CGGAATCCTG	ACGGCTCCGA	GCGGCCATGG	TGCTACACTA	CGGATCCGCA	GATCGAGCGA	780
GAGTTCTGTG	ACCTCCCCCG	CTGCGGGTCC	GAGGCACAGC	CCCGCCTCGA	GGGCGGTGGC	840
GGTTCTGGTG	GCGGTGGCTC	CGGCGGTGGC	GGTTCTCTAG	AGGGACAAAG	GAAAAGAAGA	900
AATACAATTC	ATGAATTCAA	AAAATCAGCA	AAGACTACCC	TAATCAAAAT	AGATCCAGCA	960

CTGAAGATAA AAACCAAAAA AGTGAATACT GCAGACCAAT GTGCTAATAG ATGTACTAGG 1020
AATAAAGGAC TTCCATTCAC TTGCAAGGCT TTTGTTTTTG ATAAAGCAAG AAAACAATGC 1080
CTCTGGTTCC CCTTCAATAG CATGTCAAGT GGAGTGAAAA AAGAATTTGG CCATGAATTT 1140
GACCTCTATG AAAACAAAGA CTACATTAGA AACTGCATCA TTGGTAAAGG ACGCAGCTAC 1200
AAGGGAACAG TATCTATCAC TAAGAGTGGC ATCAAATGTC AGCCCTGGAG TTCCATGATA 1260
CCACACGAAC ACAGCTATCG GGTAAAGAC CTACAGGAAA ACTACTGTCG AAATCCTCGA 1320
GGGAAGAAG GGGGACCCTG GTGTTTCACA AGCAATCCAG AGGTACGCTA CGAAGTCTGT 1380
GACATTCCTC AGTGTTTACA AGTTGAATGC ATGACCTGCA ATGGGGAGAG TTATCGAGGT 1440
CTCATGGATC ATACAGAATC AGGCAAGATT TGTCAGCGCT GGGATCATCA GACACCACAC 1500
CGGCACAAAT TCTTGCTGA AAGATATCCC GACAAGGGCT TTGATGATAA TTATTGCCGC 1560
AATCCCGATG GCCAGCCGAG GCCATGGTGC TATACTCTTG ACCCTCACAC CCGCTGGGAG 1620
TACTGTGCAA TTAAACATG CGCTGACAAA GCTGACGACG ACGACAAACA CCACCACCAC 1680
CACCACCACT AG 1692

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 563 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Gly Trp Leu Pro Leu Leu Leu Leu Thr Gln Cys Leu Gly Val
1 5 10 15

Pro Gly Gln Arg Ser Pro Leu Asn Asp Phe Gln Val Leu Arg Gly Thr
20 25 30

Glu Leu Gln His Leu Leu His Ala Val Val Pro Gly Pro Trp Gln Glu
35 40 45

Asp Val Ala Asp Ala Glu Glu Cys Ala Gly Arg Cys Gly Pro Leu Met
50 55 60

Asp Cys Arg Ala Phe His Tyr Asn Val Ser Ser His Gly Cys Gln Leu
65 70 75 80

Leu Pro Trp Thr Gln His Ser Pro His Thr Arg Leu Arg Arg Ser Gly
 85 90 95
 Arg Cys Asp Leu Phe Gln Lys Lys Asp Tyr Val Arg Thr Cys Ile Met
 100 105 110
 Asn Asn Gly Val Gly Tyr Arg Gly Thr Met Ala Thr Thr Val Gly Gly
 115 120 125
 Leu Pro Cys Gln Ala Trp Ser His Lys Phe Pro Asn Asp His Lys Tyr
 130 135 140
 Thr Pro Thr Leu Arg Asn Gly Leu Glu Glu Asn Phe Cys Arg Asn Pro
 145 150 155 160
 Asp Gly Asp Pro Gly Gly Pro Trp Cys Tyr Thr Thr Asp Pro Ala Val
 165 170 175
 Arg Phe Gln Ser Cys Gly Ile Lys Ser Cys Arg Glu Ala Ala Cys Val
 180 185 190
 Trp Cys Asn Gly Glu Glu Tyr Arg Gly Ala Val Asp Arg Thr Glu Ser
 195 200 205
 Gly Arg Glu Cys Gln Arg Trp Asp Leu Gln His Pro His Gln His Pro
 210 215 220
 Phe Glu Pro Gly Lys Phe Leu Asp Gln Gly Leu Asp Asp Asn Tyr Cys
 225 230 235 240
 Arg Asn Pro Asp Gly Ser Glu Arg Pro Trp Cys Tyr Thr Thr Asp Pro
 245 250 255
 Gln Ile Glu Arg Glu Phe Cys Asp Leu Pro Arg Cys Gly Ser Glu Ala
 260 265 270
 Gln Pro Arg Leu Glu Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
 275 280 285
 Gly Gly Gly Ser Leu Glu Gly Gln Arg Lys Arg Arg Asn Thr Ile His
 290 295 300
 Glu Phe Lys Lys Ser Ala Lys Thr Thr Leu Ile Lys Ile Asp Pro Ala
 305 310 315 320
 Leu Lys Ile Lys Thr Lys Lys Val Asn Thr Ala Asp Gln Cys Ala Asn
 325 330 335
 Arg Cys Thr Arg Asn Lys Gly Leu Pro Phe Thr Cys Lys Ala Phe Val
 340 345 350
 Phe Asp Lys Ala Arg Lys Gln Cys Leu Trp Phe Pro Phe Asn Ser Met
 355 360 365
 Ser Ser Gly Val Lys Lys Glu Phe Gly His Glu Phe Asp Leu Tyr Glu
 370 375 380

sub
 B1
 Cont'd

11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

Pub
Bill
cont'd

THE **NEW** **YORK** **PUBLIC** **LIBRARY**